RESEARCH METHODOLOGY

This chapter deals with the discussion on data base and research methodology adopted to achieve the objectives of present study. The chapter is divided into two parts. Part-I summarizes the research objectives and contains the information regarding the sample used for the study while Part-II deals with the statistical tools and techniques applied for data analysis along with limitations of the study.

PART- I

3.1 RESEARCH OBJECTIVES

The present study focuses on evaluating the profitability of technical analysis in Indian context. Technical Analysis is a broader term and includes various tools like patterns, indicators, elliott wave theory, pivot points etc. Among various tools, the study concentrates on few patterns, whose profitability is examined in the following objectives.

1. To investigate whether the use of head and shoulders pattern can generate superior returns in select Indian stocks.
2. To test the profitability of candlestick reversal patterns in nifty futures and select Indian stocks.
3. To suggest strategic combinations for enhancing profitability in stock trading.

These objectives are explained and analyzed in the specific chapters.

3.2 SAMPLE OF THE STUDY

The study of technical analysis is applicable to stocks, indices, commodities, futures or any tradable instrument where the price is influenced by the forces of supply and demand and liquidity is high. In Indian stock markets, stocks prices are determined by the demand and supply forces and orders are matched on the basis of price time priority, but liquidity is major issue. All the stocks are not highly liquid as their impact cost is too high. So, it is utmost important to select liquid stocks for this study. The CNX Nifty Index is the flagship index of NSE (National Stock Exchange, India) that comprises of 50 stocks accounting from 13 sectors of the economy. It represents about
65% of the free float market capitalization of the stocks listed on NSE, whose total traded value is approximately 46% of the traded value of all stocks on the NSE as on March 31, 2016\(^1\). One of the criteria of selecting the component stocks is liquidity. In order to qualify for inclusion in the index, it is mandatory that the security should have traded at an average impact cost of 0.50% or less during the last six months for 90% of the observations for a basket size of Rs.20 million\(^2\). Hence, it was decided to choose the stocks comprising CNX Nifty Index as the sample of study as they are highly liquid while liquidity is not a criterion for other indices.

The profitability of head and shoulders pattern and candlestick reversal patterns is examined using daily opening, high, low and closing prices of one month nifty futures contract and stocks comprising CNX Nifty Index from January 1, 2000 till March 31, 2014, covering two bull and two bear periods. The one month nifty futures contract is used here to study the profitability of patterns made on the daily data of CNX Nifty index because it is tradeable like shares and both long and short position can be made on it while the same is not possible with the spot values of CNX Nifty index. When a stock is replaced in CNX Nifty, it is also replaced in this study. So, a stock which is actually in CNX Nifty is used for analysis in this study. During this period, there were fifty one changes made in Nifty 50. Hence, the daily data of ninety seven stocks along with nifty futures is used to achieve first two objectives.

The third objective aims at suggesting strategic combinations for enhancing profitability in stock market. For this purpose, all the stock indices at NSE are selected. Firstly, the profitability of all indices is evaluated and then, head and shoulders top pattern is incorporated to check if it enhances the profits of passive investors. Till March 31, 2014, there were 36 stock indices at NSE. They are CNX Nifty, CNX Nifty Junior, LX 15, CNX 100, CNX 200, CNX 500, CNX Midcap, Nifty Midcap 50, CNX Smallcap, CNX Auto, CNX Bank, CNX Energy, CNX Finance, CNX FMCG, CNX IT, CNX Media, CNX Metal, CNX Pharma, CNX PSU Bank, CNX Realty, CNX Commodities, CNX Consumption, CNX Infra, CNX MNC , CNX PSE, CNX Service sector, CNX Nifty SHARIAH, CNX 500 SHARIAH, CNX100 Equal Weight, CNX

\(^1\)https://nseindia.com/products/content/equities/indices/nifty_50.htm accessed on August 31, 2016.

Defty, CNX Dividend opp, CNX Alpha, CNX High Beta, CNX Low Volatility, CNX NV20, and CNX NI15.

3.3 SOURCES OF INFORMATION

The daily data of opening, high, low and closing prices of one month nifty futures contract and stocks comprising CNX Nifty Index is sourced from the website of national stock exchange, India and adjusted for corporate actions like bonus, split and rights issue. The profitability of indices is evaluated using total return index values (TRI). The TRI values reflect the returns arising from dividend receipts and price movement of the constituent stocks. The TRI values of all indices are taken from bloomberg database. The implicit yield of 91-day treasury bills of Government of India is taken as a proxy of risk free rate (Connors and Sehgal, 2001). The data of 91 day treasury bills is taken from the website of Reserve Bank of India (www.rbi.gov.in). CNX Nifty is taken as market proxy. The remaining required data w.r.t to size, value and momentum factors for Indian markets was obtained from the website of IIM Ahmedabad3(Agarwalla et al., 2013).

PART-II

3.4 CALCULATION OF MEASURES OF RETURNS AND RISK

3.4.1 Calculating Continuous compounded Returns

In order to measure the profitability of the indices using total return index values, the continuously compounded return for that particular period are calculated using the following formula.

\[ R_t = LN \left[ \frac{P_t}{P_{t-1}} \right] \]  

(1)

Where,

- \( R_t \) refers to continuously compounded return
- LN refers to the natural logarithm

P_t is the price of the stock at time ‘t’.

P_{t-1} is the price of the same stock at time ‘t-1’.

However, total return values of the stock prices are not available and it is important to include dividends while measuring the profitability of the pattern. As a result, the dividends are added in the above formula. The new formula which includes dividends is as follows:

\[ R_t = LN \left( \frac{P_t + DPS_t}{P_{t-1}} \right) \]  

(2)

Where,

R_t refers continuously compounded total returns

LN refers to the natural logarithm

P_t is the price of the stock at time ‘t’.

P_{t-1} is the price of the same stock at time ‘t-1’.

DPS is the dividend per share at time ‘t’.

3.4.2 Annualized return

The annualized return refers to the annual compounded return earned by an investor over a period by investing in an asset. It is useful in a way that for comparing returns earned over different lengths of time, the returns are rescaled to one year. It is calculated as follows:

\[ R = \left( \frac{X_t}{X_o} \right)^{\frac{1}{t}} - 1 \times 100 \]  

(III)

Where: R= Annualized Return (expressed as percentage), X_t =Terminal Value, X_o =Initial Value, t= Number of years.

3.4.3 Sharpe ratio

The Sharpe Ratio measures the risk premium return earned per unit of total risk. It is calculated by dividing the excess of average daily portfolio rate of return over
average daily risk free rate with the standard deviation of excess average daily portfolio returns. It is stated as follows:

\[ S_i = \frac{\bar{R}_i - \bar{RFR}}{\sigma(R_i)} \]  

(IV)

Where: \( S_i \) = Sharpe ratio for a portfolio, \( \bar{R}_i \) = Mean return on the portfolio, \( \bar{RFR} \) = Mean return on 91-day RBI Treasury bills (proxy for risk-free rate of interest), \( \sigma(R_i) \) = Standard deviation of the daily portfolio returns.

The Sharpe ratio shows the excess return earned by an investor for per unit of variability, they are exposed to by holding a riskier asset. A portfolio with highest positive Sharpe ratio is considered best for investment while the one having negative Sharpe ratio indicates that it failed to generate any superior return over risk free rate.

3.4.4 Annualized standard deviation

Annualized Standard Deviation is a measure of volatility. An index with high annualized standard deviation is considered more volatile and hence, more risky. It is calculated as follows:

\[ \sigma_A = \sigma_d \times \sqrt{T} \]  

(V)

Where: \( \sigma_A \) = Annualized Standard Deviation, \( \sigma_d \) = Standard Deviation computed using daily returns, \( T \) = Number of trading days in a year

3.5 STATISTICAL TECHNIQUES

The study makes use of various statistical techniques which are suitable to conduct the analysis of the data for achieving the objectives of the study. The main statistical techniques used in the study are:

- The binomial test and bootstrapped skewness adjusted t-test: They are conducted using R Statistical Software (Version 3.3.1), a language and environment for statistical computing and graphics.

- The regression analysis: It is carried using E-Views Software (Version 8), a windows based econometric software by IHS Global Inc. software solutions.

The brief description of techniques is as follows:
3.5.1 Binomial Test:

It is employed in the second objective which involves testing the profitability of candlestick reversal patterns. It is used to test the statistical significance of winning trades i.e. whether the winning trades are as frequent as losing trades or not. Under this method, the proportion of the winning trades (X) is computed, which is converted to a z-score using the probability of chance (p) and its reciprocal (q) (which are used to compute the standard deviation of the distribution) along with the total number of trades (n). The formula to calculate z score is as follows:

\[
Z = \frac{x - p}{\sqrt{pq/n}}
\]  

Where,

\( Z \) = z-statistic

\( x \) = Number of winning trades

\( p \) = Hypothesized value of winning trades, here \( p=0.5 \)

\( q \) = 1-\( p \), Hypothesized value of losing trades

\( n \) = Total number of trades

The null hypothesis tested here is:

\( H_0 : p=0.5 \) (for the frequency of winning trades).

The null hypothesis \( H_0: p=0.5 \) for the frequency of winning trades is tested here because the winning rate of a candlestick pattern is compared to that of a random system, which has 50 percent chance of producing a winning trade. If the ratio of winning trades is more than 0.5 and null hypothesis is rejected, then it shows that ratio of winning trades is superior over any random system.

3.5.2 Bootstrapped Skewness Adjusted t-test:

The conventional t-test requires data to be normally distributed, where there is no skewness and kurtosis. This study is conducted using data of stock returns. But the
stock returns data suffers from skewness (Barber and Lyon, 1997). This is because the returns on the lower side are restricted to -100% while no restriction is placed for upside as stock price can increase to any level. This results in positive skewness, leading thereby to negatively biased t-statistics. Lyon et al. (1999) conclude “This leads to an overstated significance level for lower-tailed tests (i.e., reported \( p \) values will be smaller than they should be) and a loss of power for upper-tailed tests (i.e., reported \( p \) values will be too large)”. Between the skewness and kurtosis, skewness is more serious problem than kurtosis (Sophister, 1928; Neyman and Pearson, 1928 and Nair, 1941). Hence, it becomes necessary to eliminate skewness bias. This can be done using skewness-adjusted \( t \)-statistic as proposed by Johnson, 1978. Johnson develops it by making adjustment in the conventional \( t \)-statistic using skewness. The formula for the same is as follows:

\[
t_{sa} = \sqrt{n} \left( S + \frac{1}{3} \gamma S^2 + \frac{1}{6n} \hat{\gamma} \right)
\]  

(\text{VII})

Where,

\[
S = \frac{\overline{AR}}{\sigma(\overline{AR})} \quad \text{and} \quad \hat{\gamma} = \frac{\sum_{i=1}^{n} (AR_i - \overline{AR})^3}{n \sigma(\overline{AR})^3}
\]

\( t_{sa} = \) skewness adjusted \( t \)-statistic

\( n = \) sample size

\( S = \) conventional \( t \)-statistic

\( \gamma = \) is an estimate of coefficient of skewness

\( AR = \) abnormal stock return

\( \sigma(AR) = \) standard deviation of abnormal returns

The skewness adjusted \( t \)-statistic developed by Johnson (1978) is primarily grounded on an Edgeworth Expansion. The same has been studied in detail by Hall (1992) and Sutton (1993), whereby Sutton (1993) made the suggestion that “a bootstrapped application of Johnson’s statistic should be preferred over the \( t \) test when the parent distribution is asymmetrical, because it reduces the probability of type I
error in cases where the t test has an inflated type I error rate and it is more powerful in other situations."

Now, the next step involves constructing a bootstrapped distribution of the test statistic. The author developed the bootstrap distribution of skewness adjusted t statistic by drawing 10000 resamples from the original sample of size \( n_b = n/4 \). However, resample size \( n_b = n/4 \) has been chosen as suggested by Lyon et al. (1999). The test statistic, calculated in each resample is as follows:

\[
t_{sa}^b = \sqrt{n_b} (S^b + \frac{1}{3} \gamma^b S^{b^2} + \frac{1}{6n_b} \gamma^b)
\]  

(VIII)

Where,
\[
S^b = \frac{AR^b - AR_t}{\sigma^b (AR_t)}
\]

and
\[
\gamma^b = \frac{\sum_{i=1}^{n_b} (AR_{it}^b - AR_t)}{n_b \sigma^b (AR_t)^3}
\]

Thus, \( t_{sa}^b \), \( S^b \) and \( \gamma^b \) are the bootstrapped resample analogues of \( t_{sa} \), \( S \) and \( \gamma \) from the original sample for the \( b = 1, ..., 10,000 \) resamples.

The null hypothesis tested here is

\[ H_0: \mu = 0 \]

The null hypothesis \( H_0: \mu = 0 \) for the average return is used to test the statistical significance of profits earned upon the formation of candlestick pattern. If the average return is positive and null hypothesis is rejected, then the pattern is successful in generating significant returns.

3.5.3 Regression analysis

Regression involves single dependent variable and one or more independent variables. It determines whether the independent variables explain a significant variation in the dependent variable or not. Simply put, it is only concerned with the nature and degree of causality explained by the independent variables. The generalized regression equation is:

\[
Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \ldots \ldots \ldots \beta_nX_n + e
\]

(IX)
where,

\[ Y = \text{Dependent variable} \]

\[ \alpha = \text{Constant term} \]

\[ X_1, \ldots, X_n = \text{Independent variables} \]

\[ \beta_1, \ldots, \beta_n = \text{Regression/Beta coefficients of } X_1, \ldots, X_n \]

\[ e = \text{Error term (the difference between the actual and predicted values of dependent variable)} \]

Before carrying out the regression analysis, the data were put under test to ensure that it satisfies certain assumptions, which are discussed in detail as follows:

**Normality**

Normality states that the error term arising from the regression model is distributed normally so that half of the values lie above the mean while the other half lie below the mean (i.e. zero mean value of error term). The essence of this assumption is to nullify the effect of error term on the dependent variable and generates reliability to the slope coefficients. For testing normality, the Jarque-Bera statistic is referred. If the residuals are normally distributed, the Jarque-Bera statistic should not be statistically significant.

**Linearity**

Linearity explains the dependent variable as a linear function of the independent variable(s). More specifically, linearity states that linear function of slope coefficients and not of independent variables, which means that slope coefficients are raised to first power only. Regression can accurately estimate the relationship between dependent and independent variables when linearity is present, and if it is violated the results of the regression analysis will under estimate or overestimate the exact relationship and increase the chance of Type I and Type II errors. For the purpose of this study, the regression model tested using capital asset pricing model (CAPM) and Carhart four factor model is linear.

**Autocorrelation**

No Autocorrelation should be present in the error terms while carrying out regression analysis. In other words, model assumes that an error term related to current
observation is not correlated to the error term of the previous observation. The Durbin-Watson Statistic is referred to check the presence of first order autocorrelation in the data. The value of Durbin-Watson statistic ranges from 0-4. If the statistic is close to 2, the data reveals no problem of autocorrelation. If its value is near to 4, then there is negative autocorrelation of first order in data but if its near to 0, then data suffers from positive autocorrelation of first order.

**Homoscedasticity**

It means that the conditional variances of each error term are constant across all the observations of dependent variable. The opposite of homoscedasticity is ‘heteroscedasticity’ when the variances of the error term are not constant. There are a number of tests available for testing the assumption of heteroscedasticity, namely, Glejser’s (1969) Test, Breusch and Pagan (1979) Test, White’s (1980) Test, etc. (Baltagi, 2008). For the present study, White’s test has been used to check heteroscedasticity. Since, heteroscedasticity and autocorrelation are common issues in case of financial data, the problem is dealt using Newey and West (1987) approach where variance-covariance matrix of residuals is corrected for autocorrelation and heteroscedasticity. The same has been done by Savin et al. (2007).

**Multicollinearity**

It states that independent variables should not correlate highly with each other. The presence of multicollinearity in the model reduces the predictive power of the independent variable to the extent it is interrelated with the other independent variable. This problem may arise in Carhart four factor model which involves four independent variables (risk factors). But it is not serious as these factor values are given by Aggarwal et al. (2013) and they state “The correlations across the factor-returns are low and are in the lines of those reported from elsewhere in the world”.

The present study uses two regression models to carry out data analysis. They are:

**3.5.3.1 Capital Asset Pricing Model (CAPM)**

The capital asset pricing model is a single factor model introduced independently by Treynor (1961, 1962), Sharpe (1964), Litner (1965) and Mossin
(1966). The model states that there is only one factor that explains the portfolio returns and it is termed as market risk premium. It is also referred as systematic risk, which suggests that the change in the value of portfolio depends upon change in overall market. The unexplained part of the model is called as unsystematic risk, which arises due to factors which are unique to the individual assets and can be eliminated through diversification. A portfolio comprising 30-40 stock is enough to minimize unsystematic risk, thereby getting exposed to systematic risk only. The equation of CAPM model is as follows:

\[
(\bar{R}_i - RFR) = \alpha_i + \beta_i(\bar{R}_m - RFR)
\]

Where: \( \alpha_i = \) Jensen Alpha, \( \bar{R}_i = \) Mean return on the portfolio, \( \bar{R}_m = \) Mean Return of CNX Nifty (proxy for market), \( RFR = \) Mean return on 91-day RBI Treasury bills (proxy for risk-free rate of interest), \( \beta_i = \) Beta of the portfolio.

The null hypotheses tested are as follows:

\[ H_0 = \] A portfolio does not generate significant excess return over CNX Nifty i.e.

\[ \alpha = 0 \]

\[ H_0 = \] There is no significant difference in the relative risk of the portfolio and CNX NIFTY i.e.

\[ \beta = 1 \]

\[ H_0 = \] The particular portfolio can be replicated by CNX Nifty, using joint hypothesis i.e.

\[ \alpha = 0 \text{ and } \beta = 1 \]

The values of Jensen alpha and beta are obtained by regressing the excess portfolio returns against excess returns of CNX Nifty Index. Jensen alpha is a risk-adjusted measure of fund managers’ performance that measures the excess return on a portfolio over the expected returns as predicted by the capital asset pricing model (CAPM). If a stock/portfolio/fund/pattern generates a better return than its beta would predict, it has a positive Jensen Alpha, and if it returns less than the amount predicted by beta, it has a negative Jensen Alpha. An investment manager yields a statistically significant positive Jensen alpha, if he has a superior stock picking or market timing.
ability in excess of the benchmark. Similarly, a portfolio whose beta is more than 1 is considered more volatile and hence, more risky than the market. On the contrary, a portfolio with beta less than 1 is considered less risky than the market. Also, the joint hypothesis $H_0: (\alpha = 0 \text{ and } \beta = 1)$ is tested to check if a portfolio can be replicated by the benchmark index. If the null hypothesis is not rejected, then investing in the benchmark index, on average, is equivalent to investing in the portfolio, without any significant difference in return or risk.

### 3.5.3.2 Carhart Four Factor Model

The Carhart Four Factor model is an improvement over CAPM model as it includes more explanatory variables. According to this model, the excess portfolio returns are explained by controlling market-wide four risk factors i.e. market, size, value and momentum. The market factor is same as that used in CAPM model. The size factor refers to the difference in the returns of the portfolio of small cap stocks and large cap stocks. The value factor is computed by finding the difference between the returns of the portfolio of stocks having high book to market ratio and low book to market ratio. The momentum factor is determined by calculating the difference between the returns of portfolio comprising winning stocks and losing stocks. CNX Nifty is taken as the market proxy. The values of remaining risk factors i.e. size, value and momentum for Indian markets is obtained from the working paper of Agarwalla et al., 2013, which can be accessed from the website of IIM Ahmedabad\(^4\). The equation of Carhart four factor model is stated as follows:

$$\bar{R}_t - \bar{R}_{FFR} = \alpha_t + \beta_t(\bar{R}_m - \bar{R}_{FFR}) + S_t(SMB) + H_t(HML) + M_t(Momentum) \quad (XI)$$

Where: $\alpha_t =$ Four factor Alpha, $\bar{R}_t =$ Mean return on the portfolio, $\bar{R}_m =$ Mean Return of CNX Nifty (proxy for market), $\bar{R}_{FFR} =$ Mean return on 91-day RBI Treasury bills (proxy for risk-free rate of interest), $\bar{R}_m - \bar{R}_{FFR} =$ excess market premium, $SMB =$ premium of smallcap stocks over large cap stocks, $HML =$ premium of high book-to-market stocks over low book-to-market stocks, $Momentum =$ premium of one year winners vs losers.

The null hypotheses tested are:

\[ H_0 = \text{A portfolio does not generate significant excess return i.e.} \]
\[ \alpha = 0 \]

\[ H_0 = \text{There is no significant difference in the relative risk of the portfolio and CNX NIFTY i.e.} \]
\[ \beta = 1 \]

\[ H_0 = \text{There is no significant impact of size factor on excess portfolio returns i.e.} \]
\[ \beta = 0 \]

\[ H_0 = \text{There is no significant impact of value factor on excess portfolio returns i.e.} \]
\[ \beta = 0 \]

\[ H_0 = \text{There is no significant impact of momentum factor on excess portfolio returns i.e.} \]
\[ \beta = 0 \]

However, on exposing excess returns to additional variables (i.e. size, value and momentum) as in Carhart model, the statistically significant Jensen alpha (obtained using CAPM model) may turn into an insignificant Carhart alpha, showing that the superior returns evidenced by statistically significant Jensen alpha were obtained as a result of these factors and not due to the manager’s skill, whereas, the fund managers are paid for their skill in generating excess alpha.

### 3.6 LIMITATIONS OF THE STUDY

The limitations of the study are detailed as follows:

The first limitation of the study is that it includes stocks comprising CNX Nifty index only. Since CNX Nifty includes highly liquid stocks and technical analysis works best on highly liquid instruments, the study was restricted to them. The study focuses on stock market only. There are other highly liquid financial instruments available which could have been studied like commodity and currency futures.

The second limitation involves time period of only 14.25 years i.e. from January 1, 2000 till March 31, 2014. Indian markets gained popularity with the launch of online
trading by NSE in 1994 and with the introduction of derivative contracts in 2000. As a result, we have a very small history while western countries have a longer historical period to study. Hence, the studies focusing western and Japanese markets cover larger time duration. Although, BSE was formed in 1875 as ‘Native shares and Stock brokers Association’ but the daily data of stocks traded at that time is not available in public domain.

The other limitation pertains to the results of the study which are subject to the inherent limits of statistical techniques adopted to analyze the data.